

# SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB# 107640

Requester's Full Name Jeffrey E. Russel Examiner # 62785 Date: 11-4-2003  
 Art Unit 1654 Phone Number 308-3975 Serial Number 101031505  
 Mail Box and Bldg Room Location \_\_\_\_\_ Results Format Preferred (circle) PAPER DISK E-MAIL  
CMI-11013/CMI-9807

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention Conjugate For Mediating Cell Compartment Or Membrane-Specific Transport

Inventors (please provide full names) \_\_\_\_\_

Earliest Priority Filing Date 7-2-2002

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search SEQ ID NO: 6 (PKKKRKV) in STN,  
 in the U.S. patent application sequence database (pending, published, and  
 issued), and in Swissprot/Genesys/PIR.

Thank you.

JER

Point of Contact  
 P. Sheppard  
 Telephone number (703) 308-4499

## STAFF USE ONLY

Searcher	Type of Search	Vendors and cost where applicable
Searcher Phone #	NA Sequence (F)	STN
Searcher Location	AA Sequence (F)	Orig. (F)
Searcher Date	Structure (F)	Orig. (F)
Searcher Priority	Bibliography	Orig. (F)
Searcher Date	Citation	Orig. (F)
Searcher Priority	Fulltext	Orig. (F)
Searcher Date	Patent Family	Orig. (F)
Searcher Priority	Other	Orig. (F)

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:22:54 ; Search time 34 Seconds  
(without alignments)  
53.128 Million cell updates/sec

Title: US-10-031-505-6  
Perfect score: 36  
Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result	Query					
No.	Score	Match	Length	DB	ID	Description
-----						

1	36	100.0	85	11	Q64141	Q64141 rattus sp.
2	36	100.0	134	11	Q64142	Q64142 rattus sp.
3	36	100.0	178	12	Q9W7W7	Q9w7w7 simian viru
4	36	100.0	259	5	Q9NBV7	Q9nbv7 schistosoma
5	36	100.0	537	12	Q98ZP7	Q98zp7 simian viru
6	36	100.0	705	12	Q9J7C9	Q9j7c9 simian viru
7	36	100.0	707	12	O56326	O56326 simian viru
8	36	100.0	708	12	Q9DH70	Q9dh70 simian viru
9	36	100.0	709	12	Q9W8G8	Q9w8g8 simian viru
10	36	100.0	711	12	Q9WND0	Q9wnd0 simian viru
11	36	100.0	711	12	Q8JMI9	Q8jmi9 simian viru
12	36	100.0	711	12	Q9W8V1	Q9w8v1 simian viru
13	36	100.0	1165	10	Q9CAW2	Q9caw2 arabidopsis
14	35	97.2	278	12	Q9YMM7	Q9ymm7 lymantria d
15	35	97.2	542	3	Q96VV0	Q96vv0 fusarium ox
16	35	97.2	542	3	Q00832	Q00832 fusarium ox
17	35	97.2	16215	5	Q9NFS3	Q9nfs3 drosophila
18	35	97.2	18074	5	Q9I7U4	Q9i7u4 drosophila
19	33	91.7	145	16	Q9LAK6	Q9lak6 corynebacte
20	33	91.7	177	5	Q8T8R9	Q8t8r9 drosophila
21	33	91.7	212	17	Q8TZY3	Q8tzy3 pyrococcus
22	33	91.7	213	17	O59461	O59461 pyrococcus
23	33	91.7	229	5	Q8IMW8	Q8imw8 drosophila
24	33	91.7	286	10	Q9SIU9	Q9siu9 arabidopsis
25	33	91.7	290	11	Q9WTX1	Q9wtx1 mus musculu
26	33	91.7	361	3	O13856	O13856 schizosacch
27	33	91.7	421	5	Q9VC85	Q9vc85 drosophila
28	33	91.7	470	3	Q8NIS8	Q8nis8 neurospora
29	33	91.7	504	10	Q9M0V4	Q9m0v4 arabidopsis
30	33	91.7	662	11	Q91WW6	Q91ww6 mus musculu
31	33	91.7	672	5	O44189	O44189 caenorhabdi
32	33	91.7	688	12	Q91DF4	Q91df4 polyomaviru
33	33	91.7	688	12	Q8JUC5	Q8juc5 polyomaviru
34	33	91.7	688	12	Q8JUD5	Q8jud5 polyomaviru
35	33	91.7	688	12	Q8JJJ0	Q8jjj0 polyomaviru
36	33	91.7	688	12	Q9WM07	Q9wm07 polyomaviru
37	33	91.7	688	12	Q8V6G6	Q8v6g6 polyomaviru
38	33	91.7	688	12	Q8JWJ4	Q8jwj4 polyomaviru
39	33	91.7	688	12	Q9WM08	Q9wm08 polyomaviru
40	33	91.7	688	12	Q91DF8	Q91df8 polyomaviru
41	33	91.7	688	12	Q8JUD8	Q8jud8 polyomaviru
42	33	91.7	688	12	P88898	P88898 polyomaviru
43	33	91.7	688	12	Q8JUE2	Q8jue2 polyomaviru
44	33	91.7	688	12	Q91NL8	Q91nl8 polyomaviru
45	33	91.7	688	12	Q91DF5	Q91df5 polyomaviru

#### ALIGNMENTS

##### RESULT 1

Q64141

ID Q64141 PRELIMINARY; PRT; 85 AA.

AC Q64141;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE T2 antigen (Fragment).  
 GN T2 ANTIGEN.  
 OS Rattus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10118;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95347348; PubMed=7542587;  
 RA Eul J., Graessmann M., Graessmann A.;  
 RT "Experimental evidence for RNA trans-splicing in mammalian cells.";  
 RL EMBO J. 14:3226-3235(1995).  
 DR EMBL; S79053; AAB34943.1; -.  
 DR InterPro; IPR003133; T\_Ag\_DNA\_bind.  
 DR Pfam; PF02217; T\_Ag\_DNA\_bind; 1.  
 FT NON\_TER 85 85  
 SQ SEQUENCE 85 AA; 9740 MW; 5E356405DD647198 CRC64;

Query Match 100.0%; Score 36; DB 11; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
 |||||  
 Db 18 PKKKRKV 24

Search completed: November 12, 2003, 14:25:12  
 Job time : 36 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:22:54 ; Search time 11 Seconds  
(without alignments)  
29.926 Million cell updates/sec

Title: US-10-031-505-6  
Perfect score: 36  
Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	36	100.0	236	1 GT6_SCHMA	P46435 schistosoma
2	36	100.0	691	1 TALA_POVBA	P14999 polyomaviru
3	36	100.0	695	1 TALA_POVBK	P03071 polyomaviru
4	36	100.0	708	1 TALA_SV40	P03070 simian viru
5	33	91.7	384	1 YUY1_CAEEL	P54073 caenorhabdi
6	33	91.7	688	1 TALA_POVJC	P03072 polyomaviru
7	33	91.7	1083	1 T2D3_HUMAN	O00268 homo sapien
8	33	91.7	1872	1 T2D1_HUMAN	P21675 homo sapien
9	32	88.9	91	1 RK23_MARPO	P06390 marchantia
10	32	88.9	115	1 IF1A_PYRAB	Q9v138 pyrococcus
11	32	88.9	115	1 IF1A_PYRFU	Q8u0k5 pyrococcus
12	32	88.9	115	1 IF1A_PYRHO	O59280 pyrococcus
13	32	88.9	124	1 WNT3_EVATR	P28090 evasterias
14	32	88.9	178	1 RPOE_LISIN	Q927t3 listeria in
15	32	88.9	178	1 RPOE_LISMO	Q8y494 listeria mo
16	32	88.9	196	1 RAC2_LOTJA	Q40220 lotus japon
17	32	88.9	355	1 SUR6_MOUSE	P70279 mus musculu

18	32	88.9	361	1	SUR6_HUMAN	O75683	homo sapien
19	32	88.9	580	1	NO56_MOUSE	Q9d6z1	mus musculu
20	32	88.9	663	1	YD58_YEAST	Q03281	saccharomyc
21	32	88.9	1517	1	RPOC_CAMJE	Q9pi30	campylobact
22	32	88.9	2434	1	YCF1_OENHO	Q9mth5	oenothera h
23	31	86.1	501	1	GUAA_THEMA	Q9x2e0	thermotoga
24	31	86.1	517	1	GUAA_LACRH	O85192	lactobacill
25	31	86.1	520	1	GUAA_STRP3	Q8k7e6	streptococc
26	31	86.1	520	1	GUAA_STRPY	Q99zj7	streptococc
27	31	86.1	521	1	GUAA_NEIMA	Q9jw60	neisseria m
28	31	86.1	521	1	GUAA_NEIMB	Q9jxr2	neisseria m
29	31	86.1	522	1	KTR5_YEAST	P53966	saccharomyc
30	31	86.1	524	1	GUAA_YEAST	P38625	saccharomyc
31	31	86.1	960	1	K682_HUMAN	Q9y4c8	homo sapien
32	30	83.3	66	1	RL35_TREPA	O83821	treponema p
33	30	83.3	102	1	RK23_ODOSI	P49559	odontella s
34	30	83.3	110	1	RK23_PORPU	P51312	porphyra pu
35	30	83.3	129	1	RS11_AGRT5	Q8ue40	agrobacteri
36	30	83.3	129	1	RS11_BRUME	Q8yhl7	brucella me
37	30	83.3	129	1	RS11_RHILO	Q98n34	rhizobium l
38	30	83.3	129	1	RS11_RHIME	Q925w7	rhizobium m
39	30	83.3	130	1	RS11_THEMA	Q9xli4	thermotoga
40	30	83.3	139	1	RS12_MYCPN	P75546	mycoplasma
41	30	83.3	158	1	SSRP_BIFLO	Q8g540	bifidobacte
42	30	83.3	289	1	WR38_ARATH	Q8gwfl	arabidopsis
43	30	83.3	341	1	COA2_BFDV	P13892	budgerigar
44	30	83.3	353	1	RFC2_YEAST	P40348	saccharomyc
45	30	83.3	498	1	CIKW_DROME	P17972	drosophila

# ALIGNMENTS

## RESULT 1

### GT6\_SCHMA

ID GT6\_SCHMA STANDARD; PRT; 236 AA.

AC P46435;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glutathione S-transferase 6 (EC 2.5.1.18) (Fragment).

GN GST-6.

OS Schistosoma mansoni (Blood fluke).

OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;

OC Schistosomatoidea; Schistosomatidae; Schistosoma.

OX NCBI\_TaxID=6183;

RN [1]

RP SEQUENCE FROM N.A.

RA Vande Waa E.A., O'Leary K.A., Kessler M.A., Schuler L.A.,

RA Goes A.M., Mandal C., Doughty B.L., Tracy J.W.;

RL Submitted (XXX-1995) to the EMBL/GenBank/DDBJ databases.

CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.

CC -!- SIMILARITY: NONE TO THE GST SUPERFAMILY.

CC -!- SIMILARITY: STRONG, TO YEAST MAK16.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L06180; AAA29887.2; -.  
DR InterPro; IPR006958; Mak16.  
DR Pfam; PF04874; Mak16; 1.  
KW Transferase.  
FT NON\_TER 1 1  
SQ SEQUENCE 236 AA; 28286 MW; 55756FDDDD6085023 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 236;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
| | | | | | | |  
Db 222 PKKKRKV 228

Search completed: November 12, 2003, 14:25:35  
Job time : 12 secs

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:22:54 ; Search time 21 Seconds  
 (without alignments)  
 32.056 Million cell updates/sec

Title: US-10-031-505-6  
 Perfect score: 36  
 Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR\_76:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	36	100.0	350	2	JC7213	Cre recombinase pr
2	36	100.0	691	1	TVVPAS	large T antigen -
3	36	100.0	695	1	TVVPTB	large T antigen -
4	36	100.0	708	1	TVVPT4	large T antigen -
5	35	97.2	278	2	T30450	late expression fa
6	35	97.2	542	2	S20466	hypothetical prote
7	33	91.7	213	2	E71190	hypothetical prote
8	33	91.7	286	2	D84591	hypothetical prote
9	33	91.7	361	2	T38005	hypothetical prote
10	33	91.7	504	2	A85068	U3 snoRNP-associat
11	33	91.7	672	2	T32557	hypothetical prote
12	33	91.7	688	1	TVVPTJ	large T antigen -
13	33	91.7	746	2	T19409	hypothetical prote



14	33	91.7	754	2	B88252	protein F44G4.1 [i
15	33	91.7	1865	1	I48155	transcription init
16	33	91.7	1893	1	A40262	transcription init
17	32	88.9	91	1	R5LV23	ribosomal protein
18	32	88.9	115	2	B75179	translation initia
19	32	88.9	116	2	H71034	probable translati
20	32	88.9	132	2	S43488	homeotic protein L
21	32	88.9	137	2	PC1287	homeotic protein n
22	32	88.9	140	2	T43768	ribosomal protein
23	32	88.9	148	2	S21401	homeotic protein -
24	32	88.9	178	2	AC1770	B. subtilis RNA po
25	32	88.9	178	2	AH1394	B. subtilis RNA po
26	32	88.9	196	2	C95322	hypothetical prote
27	32	88.9	268	2	T32740	hypothetical prote
28	32	88.9	426	2	T16340	hypothetical prote
29	32	88.9	434	2	E96826	hypothetical prote
30	32	88.9	438	2	T47861	nucleolar autoanti
31	32	88.9	629	2	A69814	ABC transporter (A
32	32	88.9	663	2	S69626	hypothetical prote
33	32	88.9	813	2	E85135	hypothetical prote
34	32	88.9	1435	2	S69632	regulatory protein
35	32	88.9	1517	2	B81393	DNA-directed RNA p
36	32	88.9	2241	2	T16064	hypothetical prote
37	31	86.1	139	2	T03184	probable high-mobi
38	31	86.1	365	2	A96524	F11A17.6 [imported
39	31	86.1	367	2	T18839	hypothetical prote
40	31	86.1	395	2	T50804	hypothetical prote
41	31	86.1	501	2	G72206	GMP synthase - The
42	31	86.1	513	2	B96524	hypothetical prote
43	31	86.1	521	2	G81971	probable GMP synth
44	31	86.1	521	2	D81026	GMP synthase NMB19
45	31	86.1	522	2	S62941	probable membrane

#### ALIGNMENTS

##### RESULT 1

JC7213

Cre recombinase protein - Chinese hamster

C;Species: Cricetulus griseus (Chinese hamster)

C;Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 21-Jul-2000

C;Accession: JC7213

R;Koresawa, Y.; Miyagawa, S.; Ikawa, M.; Matsunami, K.; Yamada, M.; Shirakura, R.; Okabe, M.

J. Biochem. 127, 367-372, 2000

A;Title: Synthesis of a new Cre recombinase gene based on optimal codon usage for mammalian systems.

A;Reference number: JC7213; MUID:20198248; PMID:10731707

A;Accession: JC7213

A;Molecule type: mRNA

A;Residues: 1-350 <KOR>

A;Experimental source: ovarian tumor cells

C;Superfamily: phage P1 recombinase

Query Match 100.0%; Score 36; DB 2; Length 350;  
Best Local Similarity 100.0%; Pred. No. 23;

```
Matches      7;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;
Qy           1 PKKKKRKV 7
              |||||
Db           2 PKKKKRKV 8
```

Search completed: November 12, 2003, 14:24:24  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:22:54 ; Search time 41 Seconds  
(without alignments)  
27.100 Million cell updates/sec

Title: US-10-031-505-6  
Perfect score: 36  
Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	%		Query		DB	ID	Description
	Score	Match	Length				
1	36	100.0	7	15	AAW48253		Subcellular locali
2	36	100.0	7	16	AAW11495		SV 40 large T anti
3	36	100.0	7	16	AAW80531		SV40 derived nucle
4	36	100.0	7	16	AAW80458		SV40 virus large a
5	36	100.0	7	16	AAW76591		Nuclear localisati
6	36	100.0	7	17	AAW02625		SV40 large T antig
7	36	100.0	7	18	AAW21930		SV40 large T-antig
8	36	100.0	7	18	AAW18265		SV40 T-antigen nuc
9	36	100.0	7	18	AAW19795		SV40 virus nuclear
10	36	100.0	7	18	AAW14428		SV40 Large T antig
11	36	100.0	7	19	ABB79046		SV40 large T antig
12	36	100.0	7	19	AAW64765		SV40 nuclear local
13	36	100.0	7	19	AAW37985		SV40-antigen local
14	36	100.0	7	19	AAW56400		Nuclear localisati
15	36	100.0	7	19	AAW56418		Nuclear localisati
16	36	100.0	7	19	AAW50657		SV 40 large T anti
17	36	100.0	7	19	AAW53468		SV40 large T antig
18	36	100.0	7	19	AAW41646		Peptide used in pr
19	36	100.0	7	19	AAW44276		Human thymosin bet
20	36	100.0	7	20	AAW67195		SV40 T antigen nuc
21	36	100.0	7	20	AAW42289		Simian virus 40 (S
22	36	100.0	7	20	AAW27394		SV40 large T antig
23	36	100.0	7	20	AAW27401		SV40 nuclear local
24	36	100.0	7	20	AAW33115		Huntingdon's chore
25	36	100.0	7	20	AAW24432		SV40 nuclear local
26	36	100.0	7	20	AAW08408		SV40 nuclear local
27	36	100.0	7	20	AAW96350		Nucleus localisati
28	36	100.0	7	20	AAW92809		EP-893493 Seq ID 3
29	36	100.0	7	20	AAW95380		Nuclear localisati
30	36	100.0	7	21	AAW30402		Nuclear localisati
31	36	100.0	7	21	AAW14267		Nuclear localisati
32	36	100.0	7	21	AAW15405		SV40 large T antig
33	36	100.0	7	21	AAW10421		SV40 large T antig
34	36	100.0	7	21	AAW88067		SV40 large T antig
35	36	100.0	7	21	AAW90961		SV40 large T cell
36	36	100.0	7	21	AAW90468		SV40 large T antig
37	36	100.0	7	21	AAW81405		SV40 T antigen nuc
38	36	100.0	7	21	AAW58921		SV40 large T antig
39	36	100.0	7	21	AAW54042		Nuclear localisati
40	36	100.0	7	21	AAW55821		SV40 large T antig
41	36	100.0	7	21	AAW43812		SV40 large T antig
42	36	100.0	7	22	ABG99123		SV40 large T antig
43	36	100.0	7	22	AAW52169		SV40 nuclear local
44	36	100.0	7	22	AAW08726		SV40 large T antig
45	36	100.0	7	22	AAW05119		HIV-1 integrase nu

# ALIGNMENTS

RESULT 1

AAR48253

ID AAR48253 standard; peptide; 7 AA.

XX

AC AAR48253;

XX

DT 25-MAR-2003 (updated)

DT 29-JUL-1994 (first entry)

XX

DE Subcellular localisation signal PKKKRKV.

XX

KW Single chain antibody; sFv; heavy chain; light chain; kappa;

KW variable domain; hydrophilic linker; antibodies; organelle;

KW endoplasmic reticulum retention; nuclear retention signal.

XX

OS Synthetic.

XX

PN WO9402610-A1.

XX

PD 03-FEB-1994.

XX

PF 16-JUL-1993; 93WO-US06735.

XX

PR 17-JUL-1992; 92US-0916939.

PR 17-MAR-1993; 93US-0045274.

XX

PA (DAND ) DANA FARBER CANCER INST INC.

XX

PI Haseltine WA, Marasco WA;

XX

DR WPI; 1994-048868/06.

XX

PT Intracellular binding of antigens - by using antibody targetting

PT with vector system, for e.g. tumour suppression

XX

PS Claim 32; Page 101; 155pp; English.

XX

CC New vector systems comprise a sequence adapted for intracellular  
 CC delivery and expression contg. a promoter operably linked to an  
 CC antibody gene encoding an antibody which binds to a specific target  
 CC antigen. The antibody is esp. a single chain antibody in which the  
 CC heavy and light chain variable regions are joined via a hydrophilic  
 CC linker peptide. Localisation sequences are pref. included in the  
 CC constructs. See AAR48246-9 and AAR48252-3 for pref. (known) endoplasmic  
 CC reticulum retention signals. N.B. the sequence AAR48253 is described  
 CC as an ER retention signal in the Claims but as a nuclear retention  
 CC signal in the disclosure.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 15; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKKKRKV 7  
 |||||

Db 1 PKKKRKV 7

RESULT 2

AAW11495

ID AAW11495 standard; peptide; 7 AA.

XX

AC AAW11495;

XX

DT 02-APR-1997 (first entry)

XX

DE SV 40 large T antigen nuclear localisation sequence.

XX

KW Simian virus 40; SV40; large T antigen; presentation; eukaryotic;

KW nuclear localisation sequence; cell nucleus; polynucleotide;

KW gramicidin S; gene therapy; delivery; high; transfection;

KW efficiency.

XX

OS Rhesus macaque polyoma virus.

XX

PN WO9502397-A1.

XX

PD 26-JAN-1995.

XX

PF 14-JUL-1994; 94WO-US07916.

XX

PR 14-JUL-1993; 93US-0092200.

XX

PA (REGC ) UNIV CALIFORNIA.

XX

PI Haensler J, Szoka FC;

XX

DR WPI; 1995-067149/09.

XX

PT Compsns. for introducing polynucleotides into cells - comprising a

PT dendrimer polycation operatively coupled to the polynucleotide,

PT useful e.g. for gene therapy

XX

PS Disclosure; Page 28; 89pp; English.

XX

CC The present peptide is a simian virus 40 large T antigen derived  
CC nuclear localisation sequence (NLS), which can be used to present a  
CC polynucleotide (PN) to an eukaryotic cell nucleus. A NLS/PN  
CC composition, which may further comprise a membrane permeabilising  
CC agent (e.g. gramicidin S), an intercalating agent and a DNA masking  
CC agent, is useful in gene therapy and delivers the PN with high  
CC transfection efficiency.

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 16; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7

|||||||

Db 1 PKKKRKV 7

RESULT 3

AAR80531

ID AAR80531 standard; Peptide; 7 AA.

XX

AC AAR80531;

XX

DT 02-APR-1996 (first entry)

XX

DE SV40 derived nuclear localisation sequence #1.

XX

KW Moloney murine leukemia virus; gene therapy; retrovirus vector;  
KW nuclear localisation signal sequence.

XX

OS SV40.

XX

PN WO9519368-A1.

XX

PD 20-JUL-1995.

XX

PF 12-JAN-1995; 95WO-US00129.

XX

PR 14-JAN-1994; 94US-0181335.

XX

PA (ALEX-) ALEXION PHARM INC.

XX

PI Fidel SA, Kennedy SP, Mason JM;

XX

DR WPI; 1995-263825/34.

XX

PT Retroviral vector particles comprising an onco-retroviral gag capsid  
PT protein - are used to transduce non-proliferating cells.

XX

PS Disclosure; Page 29; 61pp; English.

XX

CC This sequence represents the nuclear localisation signal from the  
CC SV40 large T antigen. This peptide may be used in the production of  
CC a mutated MoMLV oncoretroviral gag capsid protein. Retrovirus vectors  
CC comprising the mutated gag capsid protein can be used to transfer  
CC exogenous genes, preferably encoding therapeutic proteins, to non-  
CC proliferating cells including stem cells and neurons. The presence  
CC of the NLS allows the exogenous gene to enter into the nucleus of the  
CC target cell, thus allowing integration of the gene into the target cell  
CC genome. The retroviral vector can be used for ex vivo gene therapy of  
CC a broad range of inherited and acquired diseases.

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 16; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7

||||||

Db 1 PKKKRKV 7

RESULT 14

AAW56400

ID AAW56400 standard; peptide; 7 AA.

XX

AC AAW56400;

XX

DT 05-AUG-1998 (first entry)

XX

DE Nuclear localisation signal (NLS).

XX

KW Signal peptide; nuclear localisation signal; NLS;

KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;

KW nuclear translocation; treatment; immune disorder; autoimmune disease;

KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;

KW tumour growth suppressor.

XX

OS Unidentified.

XX

PN WO9811907-A.

XX

PD 26-MAR-1998.

XX

PF 15-SEP-1997; 97WO-US16217.

XX

PR 12-SEP-1997; 97US-0928958.

PR 20-SEP-1996; 96US-0026978.

XX

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX

PI Blake J, Cleaveland JS, Haffar OK, Nadler SG;

XX

DR WPI; 1998-217028/19.

XX

PT Nuclear translocation inhibitor polypeptides - comprising signal

PT sequence for delivery through the cytoplasmic membrane and at least

PT 2 nuclear localisation sequences

XX

PS Claim 10; Page 44; 69pp; English.

XX

CC Peptides AAW56400-13 represent nuclear localisation signals (NLSs). The  
CC sequence is used to construct the nuclear translocation inhibitor

CC polypeptide of the invention. Nuclear translocation inhibitor

CC polypeptides comprise a signal sequence peptide capable of delivering

CC the polypeptide through the cytoplasmic membrane into a cell, and at

CC least 2 NLSs. The polypeptides can be used to inhibit nuclear

CC translocation of a cellular protein. In addition, since the nuclear

CC translocation of certain cellular peptides is required for the host

CC organism to mount an immune response, the polypeptide inhibitors are

CC useful as immunosuppression agents. The polypeptides can therefore be

CC used for the treatment of immune disorders including autoimmune

CC diseases. The polypeptides can also be used for treating physical

CC symptoms manifested by responses to allergens which can initiate a state

CC of hypersensitivity, for the treatment of sepsis and in the prevention

CC of septic shock, antiviral agents, tumour growth suppressors, and for

CC transcriptionally modulating the expression of cellular genes.



XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 19; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.3e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7

|||||||

Db 1 PKKKRKV 7

RESULT 40

AA55821

ID AA55821 standard; peptide; 7 AA.

XX

AC AA55821;

XX

DT 06-MAR-2000 (first entry)

XX

DE SV40 large T antigen NLS sequence.

XX

KW Cellular protein; nuclear translocation; nuclear localization signal;  
KW immunosuppressant; immune response; viral infection; immune disorder;  
KW rheumatoid arthritis; multiple sclerosis; juvenile-onset diabetes; NLS;  
KW systemic lupus erythematosus; SLE; asthma; sepsis; tumor growth; SV40;  
KW fibroblast growth factor; large T antigen.

XX

OS Rhesus macaque polyoma virus.

XX

PN W09957138-A1.

XX

PD 11-NOV-1999.

XX

PF 26-APR-1999; 99WO-US08984.

XX

PR 04-MAY-1998; 98US-0072429.

XX

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX

PI Nadler SG;

XX

DR WPI; 2000-062141/05.

XX

PT Novel composition comprising a nuclear localization signal and an  
PT inhibitor of nuclear translocation, useful for preventing transplanted  
PT organ or tissue rejection -

XX

PS Claim 8; Page 42; 68pp; English.

XX

CC The invention provides a composition comprising an isolated polypeptide  
CC (I) (polypeptide inhibitor of a cellular protein nuclear translocation)  
CC which comprises a signal sequence peptide capable of delivering (I)  
CC through a cytoplasmic membrane into the cell; and a nuclear localization  
CC signal sequence (NLS) present along with an immunosuppressant. (I) or  
CC its derivatives provide useful tools for introducing an exogenous  
CC polypeptide comprising an NLS into an intact cell to inhibit nuclear

CC translocation of a cellular protein, for studying the role of nuclear  
CC translocation in the regulation of cellular processes. See AAY55812 for  
CC detailed uses of (I) and compositions containing (I). The present  
CC sequence represents a NLS sequence from SV40 large T antigen.

XX

SQ Sequence 7 AA;

Query Match 100.0%; Score 36; DB 21; Length 7;

Best Local Similarity 100.0%; Pred. No. 9.3e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7

|||||||

Db 1 PKKKRKV 7

Search completed: November 12, 2003, 14:23:50

Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:22:59 ; Search time 22 Seconds  
(without alignments)  
13.463 Million cell updates/sec

Title: US-10-031-505-6  
Perfect score: 36  
Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		8					Description
Result	Query						
No.	Score	Match	Length	DB	ID		
1	36	100.0	7	1	US-08-347-792-14	Sequence 14, Appl	
2	36	100.0	7	1	US-08-182-612B-1	Sequence 1, Appli	
3	36	100.0	7	1	US-08-181-335B-1	Sequence 1, Appli	
4	36	100.0	7	1	US-08-480-463-1	Sequence 1, Appli	
5	36	100.0	7	1	US-08-664-856A-4	Sequence 4, Appli	
6	36	100.0	7	1	US-08-240-514-3	Sequence 3, Appli	
7	36	100.0	7	1	US-08-486-421-33	Sequence 33, Appl	
8	36	100.0	7	1	US-08-801-796-4	Sequence 4, Appli	
9	36	100.0	7	1	US-08-431-357-14	Sequence 14, Appl	
10	36	100.0	7	1	US-08-470-911-33	Sequence 33, Appl	
11	36	100.0	7	2	US-08-612-302A-3	Sequence 3, Appli	

12	36	100.0	7	2	US-08-931-877-4	Sequence 4, Appli
13	36	100.0	7	2	US-08-373-190-22	Sequence 22, Appl
14	36	100.0	7	2	US-08-281-423-13	Sequence 13, Appl
15	36	100.0	7	2	US-08-486-809-33	Sequence 33, Appl
16	36	100.0	7	2	US-08-928-958-10	Sequence 10, Appl
17	36	100.0	7	2	US-08-480-445A-1	Sequence 1, Appli
18	36	100.0	7	2	US-09-072-429-10	Sequence 10, Appl
19	36	100.0	7	2	US-08-438-190A-22	Sequence 22, Appl
20	36	100.0	7	2	US-08-491-988-1	Sequence 1, Appli
21	36	100.0	7	2	US-08-822-701-12	Sequence 12, Appl
22	36	100.0	7	2	US-08-480-446-1	Sequence 1, Appli
23	36	100.0	7	3	US-08-350-215-22	Sequence 22, Appl
24	36	100.0	7	3	US-08-487-799-82	Sequence 82, Appl
25	36	100.0	7	3	US-09-139-762A-125	Sequence 125, App
26	36	100.0	7	3	US-08-920-610-8	Sequence 8, Appli
27	36	100.0	7	3	US-09-069-484-4	Sequence 4, Appli
28	36	100.0	7	3	US-09-039-555B-9	Sequence 9, Appli
29	36	100.0	7	3	US-08-874-825-20	Sequence 20, Appl
30	36	100.0	7	3	US-08-935-855-12	Sequence 12, Appl
31	36	100.0	7	3	US-09-287-145A-22	Sequence 22, Appl
32	36	100.0	7	3	US-08-663-824-20	Sequence 20, Appl
33	36	100.0	7	3	US-09-007-020-1	Sequence 1, Appli
34	36	100.0	7	3	US-08-469-433B-1	Sequence 1, Appli
35	36	100.0	7	3	US-09-140-149-22	Sequence 22, Appl
36	36	100.0	7	3	US-08-789-333F-7	Sequence 7, Appli
37	36	100.0	7	3	US-09-169-015-17	Sequence 17, Appl
38	36	100.0	7	3	US-09-085-305-30	Sequence 30, Appl
39	36	100.0	7	3	US-09-242-131A-4	Sequence 4, Appli
40	36	100.0	7	3	US-08-850-744-11	Sequence 11, Appl
41	36	100.0	7	3	US-09-133-944-7	Sequence 7, Appli
42	36	100.0	7	4	US-09-615-283-4	Sequence 4, Appli
43	36	100.0	7	4	US-08-469-899A-1	Sequence 1, Appli
44	36	100.0	7	4	US-09-369-744-4	Sequence 4, Appli
45	36	100.0	7	4	US-09-411-706-3	Sequence 3, Appli

#### ALIGNMENTS

```

RESULT 1
US-08-347-792-14
; Sequence 14, Application US/08347792
; Patent No. 5573925
; GENERAL INFORMATION:
;   APPLICANT:  Halazonetis, Thanos D.
;   TITLE OF INVENTION:  p53 Proteins With Altered
;   TITLE OF INVENTION:  Tetramerization Domains
;   NUMBER OF SEQUENCES:  37
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Howson and Howson
;     STREET:    Spring House Corporate Cntr., PO Box 457
;     CITY:     Spring House
;     STATE:    Pennsylvania
;     COUNTRY:  USA
;     ZIP:     19477
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk

```

```

;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/347,792
;      FILING DATE:
;      CLASSIFICATION:  530
;      ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST58USA
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9206
;      TELEFAX:  215-540-5818
;      INFORMATION FOR SEQ ID NO:  14:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH:  7 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;      MOLECULE TYPE:  peptide
US-08-347-792-14

```

```

Query Match          100.0%;  Score 36;  DB 1;  Length 7;
Best Local Similarity 100.0%;  Pred. No. 2.5e+05;
Matches      7;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 PKKKRKV 7
        |||||
Db      1 PKKKRKV 7

```

#### RESULT 4

```

US-08-480-463-1
; Sequence 1, Application US/08480463
; Patent No. 5661025
; GENERAL INFORMATION:
; APPLICANT:  SZOKA, FRANCIS C.
; APPLICANT:  HENSLER, JEAN
; TITLE OF INVENTION:  SELF-ASSEMBLING POLYNUCLEOTIDE
; TITLE OF INVENTION:  DELIVERY SYSTEM COMPRISING DENDRIMER
; TITLE OF INVENTION:  POLYCATIONS
; NUMBER OF SEQUENCES:  11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  CROSBY HEAFEY ROACH & MAY
; STREET:  1999 HARRISON ST
; CITY:  OAKLAND
; STATE:  CA
; COUNTRY:  USA
; ZIP:  94612
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  WordPerfect 5.1/ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  US/08/480,463

```

```

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,200
; FILING DATE: 14-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KOENIG, NATHAN P.
; REGISTRATION NUMBER: 38,210
; REFERENCE/DOCKET NUMBER: 13054.00110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 763-2000
; TELEFAX: (510) 273-8832
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-480-463-1

```

```

Query Match          100.0%; Score 36; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.5e+05;
Matches      7; Conservative    0; Mismatches    0; Indels      0; Gaps      0;

```

```

Qy      1 PKKKRKV 7
        |||||
Db      1 PKKKRKV 7

```

```

RESULT 6
US-08-240-514-3
; Sequence 3, Application US/08240514
; Patent No. 5670347
; GENERAL INFORMATION:
; APPLICANT: GOPAL, T. Venkat
; TITLE OF INVENTION: PEPTIDE-MEDIATED GENE TRANSFER
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/240,514
; FILING DATE: 11-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768

```

; REFERENCE/DOCKET NUMBER: 73521/102/CLIN  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-240-514-3

Query Match 100.0%; Score 36; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
| | | | | | |  
Db 1 PKKKRKV 7

RESULT 19

US-08-438-190A-22

; Sequence 22, Application US/08438190A

; Patent No. 5965371

; GENERAL INFORMATION:

; APPLICANT: MARASCO, WAYNE

; APPLICANT: HASELTINE, WILLIAM

; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF

; TITLE OF INVENTION: PROTEINS

; NUMBER OF SEQUENCES: 78

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

; ADDRESSEE: CUSHMAN

; STREET: 130 WATER STREET

; CITY: BOSTON

; STATE: MASSACHUSETTS

; COUNTRY: US

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/438,190A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: EISENSTEIN, RONALD I.

; REGISTRATION NUMBER: 30628

; REFERENCE/DOCKET NUMBER: 41956

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 523-3400

; TELEFAX: (617) 523-6440

; TELEX: 200291 STRE UR  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-438-190A-22

Query Match 100.0%; Score 36; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
| | | | | | |  
Db 1 PKKKRKV 7

RESULT 20

US-08-491-988-1

; Sequence 1, Application US/08491988  
; Patent No. 5973116  
; GENERAL INFORMATION:  
; APPLICANT: EPENETOS, AGAMEMNON A.  
; APPLICANT: SPOONER, ROBERT A.  
; APPLICANT: DEONARAIN, MAHENDRA  
; TITLE OF INVENTION: Compounds for targeting  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAULAY NISSEN GOLDBERG KIEL & HAND, LLP  
; STREET: 261 MADISON AVENUE  
; CITY: NEW YORK  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/491,988  
; FILING DATE: 18-DEC-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GOLDBERG, JULES E.  
; REGISTRATION NUMBER: 24,408  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-986-4090  
; TELEFAX: 212-818-9479  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO



US-08-491-988-1

Query Match 100.0%; Score 36; DB 2; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
|||||||  
Db 1 PKKKRKV 7

RESULT 23

US-08-350-215-22

; Sequence 22, Application US/08350215  
; Patent No. 6004940  
; GENERAL INFORMATION:  
; APPLICANT: MARASCO, WAYNE A.  
; APPLICANT: RICHARDSON, JENNIFER  
; TITLE OF INVENTION: INTRACELLULAR TARGETING OF ENDOGENOUS  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &  
; ADDRESSEE: CUSHMAN  
; STREET: 130 WATER STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: US  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/350,215  
; FILING DATE: 12-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EISENSTEIN, RONALD I.  
; REGISTRATION NUMBER: 30628  
; REFERENCE/DOCKET NUMBER: 41956-CP3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 523-3400  
; TELEFAX: (617) 523-6440  
; TELEX: 200291 STRE UR  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-350-215-22

Query Match 100.0%; Score 36; DB 3; Length 7;  
Best Local Similarity 100.0%; Pred. No. 2.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            1 P K K K R K V 7  
              | | | | | | |  
Db            1 P K K K R K V 7

Search completed: November 12, 2003, 14:26:09  
Job time : 22 secs

OM protein - protein search, using sw model

Run on: November 12, 2003, 14:25:15 ; Search time 29 Seconds  
 (without alignments)  
 41.457 Million cell updates/sec

Title: US-10-031-505-6  
 Perfect score: 36  
 Sequence: 1 PKKKRKV 7

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Score	Query	Match	Length	ID	Description
No.						

1	36	100.0	7	8	US-08-873-601-3	Sequence 3, Appli
2	36	100.0	7	9	US-09-157-748-10	Sequence 10, Appl
3	36	100.0	7	9	US-09-104-654-2	Sequence 2, Appli
4	36	100.0	7	9	US-09-844-813-4	Sequence 4, Appli
5	36	100.0	7	9	US-09-779-233-40	Sequence 40, Appl
6	36	100.0	7	9	US-09-119-659B-3	Sequence 3, Appli
7	36	100.0	7	10	US-09-825-012-96	Sequence 96, Appl
8	36	100.0	7	10	US-09-825-012-102	Sequence 102, App
9	36	100.0	7	10	US-09-963-206B-12	Sequence 12, Appl
10	36	100.0	7	10	US-09-916-940-7	Sequence 7, Appli
11	36	100.0	7	10	US-09-942-087A-35	Sequence 35, Appl
12	36	100.0	7	10	US-09-792-630-56	Sequence 56, Appl
13	36	100.0	7	10	US-09-966-976A-12	Sequence 12, Appl
14	36	100.0	7	10	US-09-908-153B-30	Sequence 30, Appl
15	36	100.0	7	10	US-09-874-736-4	Sequence 4, Appli
16	36	100.0	7	11	US-09-953-351-53	Sequence 53, Appl
17	36	100.0	7	11	US-09-846-033B-224	Sequence 224, App
18	36	100.0	7	11	US-09-866-512A-13	Sequence 13, Appl
19	36	100.0	7	11	US-09-911-569-2	Sequence 2, Appli
20	36	100.0	7	11	US-09-876-904A-20	Sequence 20, Appl
21	36	100.0	7	11	US-09-897-844-35	Sequence 35, Appl
22	36	100.0	7	11	US-09-915-914B-23	Sequence 23, Appl
23	36	100.0	7	11	US-09-915-914B-40	Sequence 40, Appl
24	36	100.0	7	12	US-10-177-725-136	Sequence 136, App
25	36	100.0	7	12	US-10-200-879-2	Sequence 2, Appli
26	36	100.0	7	12	US-10-226-877A-12	Sequence 12, Appl
27	36	100.0	7	12	US-10-133-973-37	Sequence 37, Appl
28	36	100.0	7	12	US-10-192-393-1	Sequence 1, Appli
29	36	100.0	7	12	US-10-245-415B-35	Sequence 35, Appl
30	36	100.0	7	12	US-10-096-550-7	Sequence 7, Appli
31	36	100.0	7	12	US-10-306-878-10	Sequence 10, Appl
32	36	100.0	7	12	US-10-312-691-8	Sequence 8, Appli
33	36	100.0	7	12	US-10-315-920-7	Sequence 7, Appli
34	36	100.0	7	12	US-10-412-105-40	Sequence 40, Appl
35	36	100.0	7	12	US-10-412-109-40	Sequence 40, Appl
36	36	100.0	7	12	US-10-407-277-1	Sequence 1, Appli
37	36	100.0	7	12	US-09-293-670-13	Sequence 13, Appl
38	36	100.0	7	12	US-10-131-616B-10	Sequence 10, Appl
39	36	100.0	7	14	US-10-115-095-2	Sequence 2, Appli
40	36	100.0	7	14	US-10-080-376-56	Sequence 56, Appl
41	36	100.0	7	14	US-10-061-395-4	Sequence 4, Appli
42	36	100.0	7	14	US-10-061-395-40	Sequence 40, Appl
43	36	100.0	7	15	US-10-170-567-2	Sequence 2, Appli
44	36	100.0	7	15	US-10-006-069A-224	Sequence 224, App
45	36	100.0	7	15	US-10-096-339-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
 US-08-873-601-3  
 ; Sequence 3, Application US/08873601  
 ; Publication No. US20020064798A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: No. US20020064798Alan, Garry P.

; APPLICANT: Payan, Donald  
; TITLE OF INVENTION: COMBINATORIAL ENZYMATIC COMPLEXES  
; FILE REFERENCE: A-63915/DJB/RMS  
; CURRENT APPLICATION NUMBER: US/08/873,601  
; CURRENT FILING DATE: 1997-06-12  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: SV 40  
US-08-873-601-3

Query Match 100.0%; Score 36; DB 8; Length 7;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
| | | | |  
Db 1 PKKKRKV 7

RESULT 7  
US-09-825-012-96  
; Sequence 96, Application US/09825012  
; Patent No. US20020122798A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Robert  
; TITLE OF INVENTION: Compounds for Targeting  
; FILE REFERENCE: 43191-256808  
; CURRENT APPLICATION NUMBER: US/09/825,012  
; CURRENT FILING DATE: 2001-04-03  
; PRIOR APPLICATION NUMBER: US 60/237,159  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: GB 0008049.9  
; PRIOR FILING DATE: 2000-04-03  
; NUMBER OF SEQ ID NOS: 102  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 96  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Nuclear localisation signal  
US-09-825-012-96

Query Match 100.0%; Score 36; DB 10; Length 7;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKKKRKV 7  
| | | | |  
Db 1 PKKKRKV 7

RESULT 19

US-09-911-569-2

; Sequence 2, Application US/09911569  
; Publication No. US20030069173A1  
; GENERAL INFORMATION:  
; APPLICANT: HAWLEY-NELSON, PAMELA  
; LAN, JIANQING  
; SHIH, POJEN  
; JESSE, JOEL A.  
; SCHIFFERLI, KEVIN P.  
; GEBEYEHU, GULILAT  
; TITLE OF INVENTION: PEPTIDE-ENHANCED TRANSFECTIONS  
; NUMBER OF SEQUENCES: 120  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GREENLEE, WINNER & SULLIVAN  
; STREET: 5370 MANHATTAN CIRCLE, SUITE 201  
; CITY: BOULDER  
; STATE: CO  
; COUNTRY: US  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/911,569  
; FILING DATE: 23-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/039,780  
; FILING DATE: 16-MAR-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SULLIVAN, SALLY A.  
; REGISTRATION NUMBER: 32,064  
; REFERENCE/DOCKET NUMBER: 32-95D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303)499-8080  
; TELEFAX: (303)499-8089  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-911-569-2

Query Match 100.0%; Score 36; DB 11; Length 7;  
Best Local Similarity 100.0%; Pred. No. 5.8e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKKKRKV 7

Db                   |||||||  
                      1 PKKKRKV 7

RESULT 21

US-09-897-844-35

; Sequence 35, Application US/09897844

; Publication No. US20030087817A1

; GENERAL INFORMATION:

; APPLICANT: Cox III, George No. US20030087817A1bert

; APPLICANT: Case, Casey Christopher

; APPLICANT: Eisenberg, Stephen P.

; APPLICANT: Jarvis, Eric Edward

; APPLICANT: Spratt, Sharon Kaye

; APPLICANT: Sangamo Biosciences, Inc.

; TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells Using

; TITLE OF INVENTION: Zinc Finger Proteins

; FILE REFERENCE: 019496-002200US

; CURRENT APPLICATION NUMBER: US/09/897,844

; CURRENT FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: 09/229,037

; PRIOR FILING DATE: 1999-01-12

; NUMBER OF SEQ ID NOS: 40

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 35

; LENGTH: 7

; TYPE: PRT

; ORGANISM: Simian virus 40

; FEATURE:

; NAME/KEY: PEPTIDE

; LOCATION: (1)..(7)

; OTHER INFORMATION: SV40 large T antigen nuclear localization sequence

; OTHER INFORMATION: (NLS)

US-09-897-844-35

Query Match                   100.0%; Score 36; DB 11; Length 7;

Best Local Similarity   100.0%; Pred. No. 5.8e+05;

Matches       7; Conservative   0; Mismatches    0; Indels       0; Gaps       0;

Qy               1 PKKKRKV 7

|||||||

Db               1 PKKKRKV 7

Search completed: November 12, 2003, 14:30:25

Job time : 29 secs